Bank of Canada Banque du Canada

Working Paper 2002-28 / Document de travail 2002-28

# Filtering for Current Analysis

by

Simon van Norden

ISSN 1192-5434

Printed in Canada on recycled paper

Bank of Canada Working Paper 2002-28

October 2002

# Filtering for Current Analysis

by

**Simon van Norden**

Ergodic Quantitative Consulting Inc.
École des Hautes Études Commerciales (HEC)
Centre Interuniversitaire de Recherche en Analyse des Organisations (CIRANO)
Centre de Recherche et Développement en Économique (CRDE)
Montréal, Quebec
simon.van-norden@hec.ca

The views expressed in this paper are those of the author.
No responsibility for them should be attributed to the Bank of Canada.

# Contents

Abstract/Résumé ..... v

1. Introduction ..... 1

2. An Optimal Band-Pass Filter for Current Analysis ..... 3

   2.1 Filters: a primer ..... 3

   2.2 Business cycle filters ..... 4

   2.3 Optimal current-analysis filters ..... 6

3. Applications ..... 8

   3.1 Data ..... 8

   3.2 Filtering productivity growth ..... 9

   3.3 Filtering inflation ..... 11

   3.4 Filtering GDP ..... 12

4. Conclusions ..... 14

References ..... 16

Tables ..... 18

Figures ..... 20

Appendix A: Derivation of the Optimal Filter ..... 24

Appendix B: Historical Antecedents ..... 33

## Abstract

This paper shows how existing band-pass filtering techniques and their extension can be applied to the common current-analysis problem of estimating current trends or cycles. These techniques give estimates that are "optimal" given the available data, so their standard errors represent a lower bound on what can be achieved with other univariate techniques. Applications to the problems of estimating current trend productivity growth, core inflation, and output gaps are considered.

*JEL classification: C1*
*Bank classification: Econometric and statistical methods; Potential output*

## Résumé

L'auteur montre comment les techniques actuelles de filtrage passe-bande et leurs prolongements peuvent servir à estimer des tendances et des cycles courants. Ces techniques donnent des estimations jugées « optimales » compte tenu des données disponibles, de sorte que les écarts-types s'y rattachant représentent la borne inférieure de la marge d'erreur qui serait associée aux résultats produits par d'autres techniques univariées. Dans cette étude, l'auteur examine les applications de ce filtre aux problèmes que posent l'estimation de la croissance de la productivité, de l'inflation fondamentale et de l'écart de production observés.

*Classification JEL : C1*
*Classification de la Banque : Méthodes économétriques et statistiques; Production potentielle*

# 1. Introduction

A common problem in macroeconomics is that of measuring the business cycle, or more generally of separating long-run trends from short-term movements. A technique that does so can be thought of as a *filter,* one that is applied to raw economic data prior to analysis. The best known is the Hodrick-Prescott (H-P) filter, which has become the benchmark against which all other filters in macroeconomics are compared.[1]

The H-P filter is unabashedly arbitrary; it was proposed and adopted largely on the basis that it gave results that "looked reasonable." Its use has since been rationalized as an approximate band-pass filter. A band-pass filter isolates movements in a series between a specified upper and lower frequency or duration; movements outside the desired frequency band are eliminated. As commonly used with quarterly economic data, the H-P filter eliminates or greatly reduces most long-run movements in the series while preserving those at roughly business cycle frequencies.[2] The result is a detrended series that looks like a business cycle and has served as an agnostic basis for much economic analysis.[3]

Given its pre-eminent role and its arbitrary nature, the H-P filter has been the focus of much emulation and innovation in recent years. Baxter and King (1999) propose replacement of the H-P filter with a more exact band-pass filter, arguing that better results come from using a better approximation. Gomez (2001) and Pollock (2000) propose the use of other ad hoc filters that are commonly used in engineering and that also approximate band-pass filters.[4] Pedersen (1998) and Kaiser and Maravall (2001) propose extensions or modifications to the H-P filter to improve its performance. Most of this literature, however, has ignored the application to current analysis.

The distinguishing feature of current analysis is that it interprets the most recently available information; the cycle (or trend) of interest is that at the end of the data sample. The dominant

---

1. Its current popularity stems from its use in the seminal working paper by Hodrick and Prescott (1977), finally published as Hodrick and Prescott (1997), although the technique dates from the 1920s.
2. In practice, business cycles are defined as having durations of between 6 and 32 quarters. This definition gained popularity after Baxter and King (1999) stated that Burns and Mitchell (1947) characterized business cycles in that way. The numbers are not written in stone; Stock and Watson (1998) quote Burns and Mitchell (1947, 3) as stating that business cycles vary in duration "from more than one year to ten or twelve years."
3. The claim that H-P-filtered output "looks like" a business cycle is a popular misstatement. Business cycle measurement and analysis, since its infancy, has used the H-P filter (or even simpler moving-average filters, which produce similar results). It is therefore probably more accurate to say that the output of these filters has *defined* what we think of as business cycles. See Morley, Nelson, and Zivot (1999).
4. Both examine Butterworth filters, of which they note that the H-P filter is simply a special case.

focus of the aforementioned literature is historical analysis, in which we care mostly about the cycle (or trend) somewhere in the middle of the sample. This distinction is sometimes critical. Most of the analysis in the existing literature is restricted to symmetric filters; to isolate the cycle or trend at time $t$, such filters use an equal number of observations from before and after $t$.[5] This precludes their use at the end of the sample. Other filters are justified on the basis of their mid-sample properties, but can behave quite differently at the end of the sample.[6]

This paper considers the filtering problem from the perspective of current analysis. Rather than consider ad hoc approximations to band-pass filters, it shows how to construct one-sided band-pass filters that are optimal in a minimum mean-squared-error (MSE) sense. Unlike the filters mentioned in the above-noted literature, the optimal filter will vary with the properties of the data series to be filtered. While such filters are little-known in macroeconomics, they are not new; Appendices A and B, respectively, review the contributions of Christiano and Fitzgerald (1999) and Koopmans (1974) and provide a modest extension of their results. The body of this paper gives an overview of optimal filters and applies them to three problems of widespread interest: estimating the current output gap, the current trend growth rate of productivity, and the current trend rate of inflation.

Literature related to that noted above examines the reliability of filtered estimates of trends and cycles.[7] Since the current analysis filter discussed in this paper minimizes an MSE criterion, we can relate its reliability to those of other filters examined in this literature. It also establishes an upper bound on the accuracy that any such filters can hope to achieve. Because this bound depends in a complicated way on the properties of the data analyzed, we investigate the properties of the optimal filter for the three common problems in current analysis.

Section 2 provides a non-technical overview of band-pass filtering and the optimal one-sided band-pass filter. Appendix B gives the derivation and a more detailed discussion of this filter. Section 3 applies the filter to three common macroeconomic problems of measuring trends and cycles and discusses the results. Section 4 summarizes the conclusions and suggests avenues for further research.

---

5. Baxter and King (1999), for example, suggest reserving about five years of data from each end of the sample to provide the necessary leads and lags.
6. The H-P filter is a case in point. Comparisons with band-pass filters are based on its symmetric MA representation, which is a limit the filter approaches in the middle of a large sample. Its representation at the end of the sample is quite different; see St-Amant and van Norden (1997).
7. Examples of this include Setterfield, Gordon, and Osberg (1992), Staiger, Stock, and Watson (1997), Orphanides and van Norden (1999, 2002), and Cayen (2001).

# 2. An Optimal Band-Pass Filter for Current Analysis

## 2.1 Filters: a primer[8]

A filter can be thought of as an algorithm for processing a time series to get a more meaningful statistic; e.g., the process of averaging measured rainfall at a given location to get "average" rainfall. We can describe this mathematically as

$$S = f(\tilde{y}), \tag{1}$$

where $S$ is our statistic, $\tilde{y}$ is our time series, and $f(.)$ is our filter. While such processing could be complex, attention often focuses on a particularly simple and tractable case: the *linear, time-invariant* filter. The distinguishing features of such filters are that the weight, $\beta_i$, we put on a particular observation does not depend on $t$, and that the operation is linear in $\tilde{y}$.[9] Such filters can therefore be described as

$$S_t = \tilde{\beta} \cdot \tilde{y} = \sum_{i=-\infty}^{\infty} \beta_i \cdot y_{t-i}. \tag{2}$$

If we think of $y_t$ as a random error term, the Wold decomposition theorem tells us that this class of filter is related to the class of autoregressive moving average (ARMA) process. In this situation, since the properties of $\tilde{y}$ are held fixed, the properties of $S$ are determined by $\tilde{\beta}$ (or vice versa).

We are often particularly interested in the dynamics of $S$, which may be uniquely characterized via frequency or spectral analysis. The idea is to decompose all the movements in $S$ into cycles of varying frequency and amplitude. Such a unique decomposition exists if $S$ is stationary (and if not, we can difference it until it is.) Furthermore, since cycles of different frequencies are uncorrelated in the long run, the variance of $S$ will simply be the sum of its variances over all frequencies. The relative importance of these different frequencies in the overall variance tells us something about the dynamic behaviour of the series. For example, an identically, independently distributed (i.i.d.) error will display a constant variance at all frequencies, while a random walk will have much more variance at low frequencies (long cycles) than at high (short cycles). The

---

8. This section provides an intuitive introduction to the filtering terms used in the rest of this paper. It can be skipped without loss of continuity by those familiar with spectral analysis.
9. Several well-known filters do not belong in this class. Examples include the H-P filter (where the weights vary with $t$) and the Hamilton filter for the probability of being in a particular regime (which is a non-linear function of $\tilde{y}$).

function decomposing the total variance by frequency is commonly called the *spectrum* or *spectral density* and is typically shown graphed from 0 (lowest frequencies, infinitely long cycles) to $\pi$ (highest observable frequencies, cycles of 2 periods).

The spectrum of $S$ depends on the properties of both $\tilde{\beta}$ and $\hat{y}$. To understand the effects of $\tilde{\beta}$, we can divide the spectrum of $S$ by that of $\hat{y}$ to define the *squared gain* or *transfer* function of $\tilde{\beta}$. Frequencies at which the squared gain is greater than 1 are accentuated in $S$, while those at which the squared gain is (close to) zero are (nearly) removed from $S$. The aim of *band-pass* filtering is to choose $\beta$ to match a particular kind of squared gain function, one that has a gain of 1 over a particular frequency range $(l, u) \mid 0 \leq l < u \leq \pi$ and zero elsewhere. The case where $l = 0$ is called a low-pass filter, while $u = \pi$ is called a high-pass filter.

If a filter has the property $\beta_j = \beta_{-j}\ \forall j$, it is called a *symmetric* filter, whereas if $\beta_j = 0\ \forall j > 0$ the filter is said to be *one-sided*.[10] Symmetric filters have the property that $S$ will tend neither to lead nor to lag movements in the corresponding components of $y$; the same is not generally true for non-symmetric filters.[11] This effect of non-symmetric filters is called *phase shift* and in general will vary from one frequency to another.

## 2.2 Business cycle filters

Band-pass filtering is an approach to the measurement of trends or cycles in macroeconomics whose appeal rests on two key assumptions:

(i) We can agree on some threshold duration such that we wish to interpret movements of longer duration as trends and those of shorter duration as cycles.

(ii) Aside from this, we wish to remain fairly agnostic about the economic or stochastic processes generating the data.

We can detrend the data using a low-pass filter (one that passes all frequencies *below* the threshold; i.e., all durations *above* the threshold) to isolate the trend, or, equivalently, use a high-pass filter to isolate the cycle. Two of the three applications we study below use low-pass filters to isolate trends. The case of business cycles is more complex, since we wish to exclude both the trend component and a seasonal/short-lived component. We therefore need a band-pass filter to block the movements both of very long and very short duration.

---

10. The only one-sided symmetric filter is the trivial filter, which just multiplies $y_t$ by a constant; we will ignore that special case and proceed as if these two classes are mutually exclusive.
11. For example, consider the difference between a centre-weighted and a one-sided moving average.

The ideal band-pass filter would have a gain of zero outside the $(l, u)$ interval and a gain of one inside. Deviations from the former condition would allow leakage from undesired frequencies, while deviations from the latter would distort the "true" cycle present in the data. This unique filter exists and is given by the formula

$$B_j = \frac{\sin ju - \sin jl}{\pi j} \text{ for } |j| \geq 1 \qquad (3)$$
$$= \frac{u - l}{\pi} \text{ for } j = 0.$$

One problem with this ideal filter is that we require the sum in equation (2) to go from $-\infty$ to $\infty$. Truncating this sum at some finite values, say $-N$ and $N$, results in an approximate filter in which the desired rectangular shape of its squared gain is contaminated by sinusoidal imperfections (see Figure 1). Baxter and King (1999) suggest that using values of $N$ as small as 20 gives reasonable results for U.S. business cycles if we adopt the Burns and Mitchell cut-offs of 6 and 32 quarters.

The problem with this approach is that it cannot be used for current analysis. Using the Baxter and King approximation with N = 20 on quarterly data would imply that our most recent estimates of the business cycle would be 20 quarters prior to the last quarter for which we had data. One way around this would be to use the Baxter-King formula at the end of the sample, simply omitting (i.e., replacing by zero) the missing observations, which are not yet known. As Figure 1 shows, this gives poor results, even for large values of N; the resulting filters have a gain that is far from 1, vary considerably over the frequency band of interest, and leak much more of the frequencies outside the desired band. Stock and Watson (1998) use a different ad hoc solution. They fit the available time series to a simple autoregression model, then use forecasts from the fitted model in place of the required future observations. Unfortunately, they do not provide a justification for this procedure, nor do they examine how closely it approximates the ideal filter.

Another criticism of this approach is that even if the filter used has a gain function close to that of the ideal filter, this need not imply that the series it produces will be a good approximation of the ideally filtered series.[12] The problem is that many economic series display a "typical Granger spectral shape": the density in their spectrum is highly concentrated at the lowest frequencies. This in turn means that for band-pass or high-pass filtering (e.g., for measuring business cycles) we care much more about how well we approximate the ideal filter at low frequencies than at high frequencies.[13]

12. For example, see Guay and St-Amant (1997).
13. Pedersen (1998) re-examines H-P filters from this perspective and suggests alternatives to the traditional value of 1600 for its smoothing parameter.

## 2.3 Optimal current-analysis filters

To adapt the band-pass filtering approach to a current analysis context, we would like to have some optimal filter, $\{\hat{B}_j\}$, which minimizes

$$E\left[\sum_{j=0}^{T-1} \hat{B}_j \cdot y_{T-j} - \sum_{j=-\infty}^{\infty} B_j \cdot y_{T-j}\right]^2. \tag{4}$$

In other words, given $T$ observations on the series we wish to filter, $\{y_t\}$, the optimal filter will give us the minimum MSE estimate of what the ideal filter would give us with data from $\infty$ to $-\infty$. As Appendix A shows, this problem has a unique solution under fairly standard conditions.[14] In the case where $\{y_t\}$ is stationary, we find that

$$\hat{\vec{\beta}} = \hat{\Sigma}_y^{-1} \cdot \vec{B} \cdot \vec{\sigma}_y, \tag{5}$$

where

$\hat{\vec{\beta}} = [\hat{B}_0, \ldots, \hat{B}_{T-1}]'$, a length $T$ column vector,

$\hat{\Sigma}_y = [\vec{\sigma}_{y,0}, \vec{\sigma}_{y,1}, \vec{\sigma}_{y,2}, \ldots, \vec{\sigma}_{y,T-1}]'$, a $T$ x $T$ matrix,

$\vec{\sigma}_{y,j} = [\sigma_y(-j), \sigma_y(-j+1), \sigma_y(-j+2), \ldots, \sigma_y(T-1-j)]'$, a length $T$ column vector,

$\vec{\sigma}_y = [\sigma_y(-q), \sigma_y(-q+1), \sigma_y(-q+2), \ldots, \sigma_y(q)]'$, a $2q+1$ column vector,

$\sigma_y(q) = cov(y_t, y_{t+q})$,

$\vec{B} = [\vec{B}^0, \vec{B}^1, \vec{B}^2, \ldots, \vec{B}^{T-1}]'$, a $T$ x $2q+1$ matrix,

$\vec{B}^j = [B_{j-q}, B_{j-q+1}, B_{j-q+2}, \ldots, B_{j+q}]'$, a $2q+1$ column vector, and

$B_j$ is the weight of the ideal filter, given by equation (3).

In general, these optimal weights depend on (i) the number of observations, $T$, we have, (ii) the dynamics of our series, $y_t$, as measured by its autocovariances, and (iii) the ideal weights, $B_j$,

14. Appendix A gives solutions for the case of integrated autoregressive moving average (ARIMA)(p,d,q) processes ($-1 < d < 3$) and surveys related contributions in the literature.

from equation (3). Furthermore, they have an intuitive interpretation as the solution to a regression problem—one where we regress the doubly infinite set $\{B_j \cdot y_{T-j}\}$ on our $T$ observations, $\hat{\gamma}$. The resulting coefficients are our optimal weights, and therefore our minimization problem, equation (4), simply seeks to minimize the variance of the regression residuals.[15]

We can better understand this formula and the above intuition if we consider two special cases. First, suppose that $y_t$ is i.i.d. This means that $\sigma_y(l) = 0$ for all $l \neq 0$, so we may set $q = 0$. This makes $\hat{\sigma}_y$ a scalar, equal to the variance of $y$, and $\hat{\Sigma}_y$ is simply the identity matrix times this variance. This means that equation (5) further simplifies to $\hat{\beta} = [B_0, \ldots, B_{T-1}]$. Put another way, in this case the optimal solution is simply a truncated version of ideal weights, precisely the same solution that was seen to give very poor results in Figure 1.

Now suppose that $y$ follows a stationary MA($Q$) process. Since its autocovariances will be zero for leads and lags greater than $Q$, this again effectively determines $q$ in equation (5). Suppose, however, that instead of using the optimal weights, we use the Stock and Watson approach of padding our $T$ observations with $Q$ forecasts/backcasts from the MA model at each end of the sample, and then use the Baxter-King approximate filter with $N = Q$. The estimate from this two-step ad hoc procedure will be identical to the estimate from our optimal filter, because the optimal weights given in equation (5) reflect both the weights used to form forecasts/backcasts at the ends of the available sample and the weights that the Baxter-King filter would place on them. Put another way, equation (5) implies that the Stock-Watson two-step procedure will give optimal estimates at the end of the sample provided that (i) we use the "right" forecasting model to pad our data, and (ii) we pad our sample until our forecasts have converged to zero.

Another feature of equation (5) is that it lets us solve for the minimum value of equation (4). This is useful, since it tells us how well our best end-of-sample estimates can approximate the ideal estimates. The general solution is given by

$$E\left[\sum_{j=-\infty}^{\infty} B_j \cdot y_{T-j}\right]^2 - E\left[\sum_{j=0}^{T-1} \hat{B}_j \cdot y_{T-j}\right]^2. \tag{6}$$

In the above case, where $y$ is i.i.d., this reduces to

15. This interpretation is developed further in Appendix A, particularly in section A.1.

$$\left(\sum_{j=-\infty}^{\infty} B_j^2 - \sum_{j=0}^{T-1} B_j^2\right)\cdot\sigma_y^2 = \left(\sum_{j=-\infty}^{-1} B_j^2 + \sum_{j=T}^{\infty} B_j^2\right)\cdot\sigma_y^2. \tag{7}$$

# 3. Applications

We now examine the performance of the optimal filter by applying it to three problems of common interest: estimating the current output gap, the current trend growth rate of productivity, and the current trend rate of inflation. The first of these differs from the other two in that we analyze the raw data in levels, not growth rates, and in that we seek to isolate the intermediate frequencies rather than the low frequencies. Because the optimal filter is a function of the dynamic properties of the series analyzed, our results can be expected to differ across applications.

## 3.1 Data

The output series ($Y_t$) is the natural logarithm of Canadian real GDP (1997 dollars) for the period 1961Q1 to 2001Q1. The inflation series ($\pi_t$) is the monthly difference of the natural logarithm of the Canadian all-items CPI (seasonally adjusted) covering the period January 1948 to August 2001. The productivity series ($Q_t$) is the natural logarithm of quarterly data on Canadian real GDP per hours of employment from 1966Q1 to 2001Q2; hours worked are average weekly hours times the number of people employed, both of which are from Statistics Canada's labour force survey. Note that deterministic components were removed from all three series prior to analysis. In the case of $Q_t$ and $\pi_t$, the series were demeaned; for $Y_t$ a deterministic linear trend was also removed.

For each series, two different estimates of the autocovariance function, $\sigma_y(q)$, were then constructed. The first fit a low-order ARMA model to the data (Table 1), and then used the estimated parameters of the ARMA model to calculate the implied covariances.[16] The second used a nonparametric kernel estimate.[17] In both cases, each data sample of $N$ observations was used to calculate $N-1$ autocovariances. The two approaches gave sometimes similar estimates, as shown in the top-left panel of Figures 2 through 4, with the nonparametric kernel tending to

16. The *ARMABIC3()* procedure from the Coint module for GAUSS, by Ouliaris and Phillips (1995), was used for estimation and model selection. This uses the BIC criterion for model selection and a 2- or 3-stage Hannan-Rissanen iterative estimation procedure.

17. The results presented here use the quadratic-spectral kernel (without the data-dependent bandwidth selection.) Limited experimentation suggested that the results were not sensitive to this choice.

capture somewhat more complex and persistent dynamics. The second panel (top right) of Figures 2 through 4 shows the corresponding spectrum for each series.

The dynamics of the three series look quite different. Quarterly productivity growth shows very little persistence and is well-approximated as white noise; its spectrum is therefore relatively flat. Output shows slowly decaying autocovariances, consistent with estimated autoregressive roots of nearly 0.95. Its spectrum displays the typical Granger shape, with density powerfully concentrated in the low frequencies. Monthly inflation falls between these two extremes, with the dynamics of its ARMA approximation showing less persistence than those of the nonparametric kernel estimate.[18] The result is two very different-looking spectra, with the ARMA-based estimate looking quite flat, but the kernel-based estimate showing even more concentration in the low frequencies than the spectrum for output growth.

## 3.2 Filtering productivity growth

Trends in productivity growth are the subject of considerable analysis, with interest in current trends having intensified in recent years. While it is widely acknowledged that labour productivity is procyclical, most analysis of productivity growth does little to explicitly separate its trend and cyclical components beyond examining averages of growth rates over several years. It would therefore be useful to construct optimal estimates of current trend productivity, as well as to know how reliable such estimates may be. This is precisely what the results from section 2 now enable us to do.

Since we are trying to remove business cycle influences from productivity growth, we adopt the Burns-Mitchell-Baxter-King characterization of these cycles as having durations of up to eight years in length. Our ideal filter for quarterly data is therefore a symmetric low-pass filter that blocks all frequencies above $\pi/16$. This, together with the results presented in section 3.1, is all we need to construct the optimal filter. Its properties are described in Figure 2 and Table 2.

The third panel in Figure 2 (middle row, left column) compares the weights of the optimal and ideal filters. In the case of the ARMA model, the two are identical, because the model approximates quarterly productivity growth as uncorrelated; this is precisely the simple case discussed in section 2.3. The weights for the kernel model are similar, but die away more slowly and are somewhat more volatile, reflecting the greater persistence in the series that the kernel detects.

18. ARMA model selection for the CPI data was problematic. It is doubtful that the MA(12) adequately captures the persistence of inflation, since this would imply that inflation shocks completely die out in 12 months.

Panels 4 and 6 in Figure 2 (right column, middle and bottom rows) compare the spectral properties of the three filters and the two optimally filtered series. The gain of the optimal ARMA filter shows the same pattern we saw previously in Figure 1 for other truncated ideal filters; the only difference is the number of observations in the approximation.[19] The optimal kernel filter has a more complicated gain function, tending to fluctuate around the gain of the optimal ARMA filter. Note that the kernel filter has a particularly high gain for some of the frequencies between 0 and 0.1, and a particularly low gain for some of the frequencies between 0.1 and 0.2. This relative emphasis is to be expected, given the peaks and valleys of the kernel-estimated spectrum in these ranges, and therefore their relative importance in the overall performance of the filter.

Multiplying the spectral density at a given frequency from panel 2 by the squared gain at that frequency from panel 4 gives us the spectrum of the filtered series shown in panel 6. In contrast to the relatively flat spectra shown in panel 2, our low-pass filters have succeeded in powerfully concentrating the spectrum of the filtered series over the desired frequencies. At the same time, however, the imperfections of the optimal filters are clearly visible; the spectral density at the low frequencies is only about half that in the raw data, and the ARMA-filtered spectrum has an artificial peak at the cut-off frequency, which implies a spurious cycle in the filtered series.[20]

The overall performance of the filter across all frequencies is summarized by the statistics presented in Table 2. The variance of the raw series (line 1) is simply the area under the spectra shown in panel 2 of Figure 2.[21] Similarly, the variance of the ideally filtered series (line 2) is the area under those spectra lying between frequencies 0 and $\pi/16$. In this case, the ideally filtered trend accounts for roughly 1/20th of the total variance of the observed series. The optimally filtered estimate of that trend captures little more than half of the variance of the ideal trend. Put another way, the MSE of the optimally filtered estimate will be almost as variable as the estimate itself; the noise-signal ratios are slightly over 0.9 and the correlation between the optimal and the ideal estimates will be only about 0.72. Not surprisingly, given the similarities of the ARMA and kernel estimates of the autocovariance functions, the two optimal filters perform similarly as well.

19. To understand why a filter with a gain everywhere less than 1 may still be optimal, consider the effect of scaling all the filter weights by some constant $k>1$. This has the effect of scaling the squared gain everywhere by $k^2$. In the case of the ARMA filter, this will reduce the difference between its gain and that of the ideal filter at frequencies below the cut-off frequency (i.e., reducing compression), thereby improving the estimate. This benefit is counterbalanced, however, by the effect of increasing the difference between the two filters at frequencies above the cut-off (i.e., increasing leakage). The optimal scale is the one at which the marginal benefits at some frequencies of a change in scale are exactly equal to the marginal costs at all other frequencies.
20. Of course, the current application is designed to produce a single point estimate, rather than a series, so this point may be moot.
21. Strictly speaking, it is twice that area, since the full spectrum is symmetric about 0; the figure shows only half that range. This applies to the analysis of subsequent rows in this table, as well.

## 3.3 Filtering inflation

As the objective of price stability has moved to the forefront of monetary policy formulation throughout much of the world over the past decade and a half, more attention has focused on how to measure and monitor progress towards this goal. Most monetary authorities propose a modified (or "core") measure of inflation that aims to capture persistent trends in inflation or inflationary pressures.[22] It would therefore be of interest to use band-pass filters to construct optimal measures of current trends in inflation, where these trends are again defined using a low-pass filter. The appropriate cut-off frequency to use for such a filter is debatable; for the example we study in this section, a frequency of $\pi/24$ (corresponding to cycles lasting 48 months) is used to give seasonal influences and short-run nominal shocks ample time to dissipate. Results are presented in Table 3 and Figure 3.

Panel 3 (left column, middle row) in Figure 3 shows us that the optimal weights are now more different from the ideal weights. While the ideal weights are quite small and vary gradually over time, both sets of optimal weights put much higher weights on the most recent observations. Again, the ARMA model gives weights that converge more quickly to the ideal weights than the kernel model, presumably reflecting the greater persistence in the kernel-estimated dynamics.

Panel 4 (right column, middle row) shows that the gain functions of the optimal filters are again different from that of the ideal filter in several respects. The gain function for the ARMA filter resembles the shape shown in Figure 1, with a gain of less than 0.5 for most of the pass band, and a gain near 1 for only a narrow band near the cut-off frequency. Although the gain drops sharply beyond that point, it stays significantly above zero for the remainder of the graphed frequency range. The gain of the kernel filter bears only a rough resemblance to that of the ARMA filter, the most striking difference being the presence of multiple narrow spikes in the gain function throughout the pass band and stop band. The gain within the pass band varies by a factor of roughly 10-to-1, and there are multiple peaks well outside the pass band with gains close to or exceeding 1.

Panel 6 (right column, bottom row) shows that, despite the apparently irregular gain functions, both of the resulting filtered series capture most of the density of the raw series at the low frequencies and have a very sharp drop in density at the cut-off frequency, with very little density at the higher frequencies. This reflects the fact that both filters have sharp drops in gain at the cut-off frequency, and that the potentially large leakage they allow from much higher frequencies is

22. Unlike the techniques examined in this paper, many approaches to measuring core inflation rely on the analysis of disaggregated price movements.

relatively unimportant, owing to their lack of importance in the original spectrum. Similarly, the variable gain of the kernel filter within the pass band does not appear to greatly distort the spectrum of inflation within this band. This is because peaks in the raw spectrum correspond to frequencies around which the gain is not far from 1, while areas of particularly low gain in the pass band correspond to troughs in the raw spectrum.

Table 3 confirms the relatively good performance of the optimal filter, particularly for the kernel model. The optimal ARMA filter performs somewhat better than its counterpart in the case of productivity growth, giving correlations with the ideal filter of almost 80 per cent and a noise-signal ratio of 58 per cent. The optimal kernel filter, however, performs much better, giving a 95 per cent correlation with the ideal filter and a noise-signal ratio of only 11 per cent.

This improved performance is consistent with the basic intuition observed in section 2.3, that persistence tends to improve the quality of current estimates of the trend. We can understand this in terms of the Stock-Watson two-step procedure: the more persistent a series is, the better able we are to forecast it and therefore the better we approximate the ideal filter. Since kernel estimates of the autocovariance function imply much longer dependence than the MA(12) model chosen by the ARMA model-specification procedure, they imply that we should be able to forecast inflation much further into the future and therefore to reduce the difference between mid-sample and end-of-sample filter estimates.

Another way to understand these results is to recall that as persistence increases, the spectrum of our data series becomes increasingly concentrated in the lowest frequencies. Since our goal is to design a low-pass filter, this in turn reduces the importance of leaking higher frequencies, allowing an increase in the average gain of the optimal filter and thereby a better approximation of the ideal filter.

## 3.4 Filtering GDP

Optimal filters for business cycles and their properties are described in Figure 4 and Table 4. Unlike the two previous filters considered, this is a band-pass rather than a low-pass filter, and it uses the values suggested by Baxter and King $(\pi/16, \pi/3)$ to define the frequencies of interest.

Panel 3 (left column, middle row) of Figure 4 shows that, aside from the first few observations, the optimal ARMA weights are indistinguishable from the ideal weights. The kernel weights, however, again display much more variation and die away more slowly. Panel 4 (right column, middle row) shows that the squared gain of the optimal ARMA filter resembles the patterns shown in Figure 1. The gain function for the kernel filter roughly follows that of the ARMA filter,

but with greater volatility, reflected in large, narrow peaks and troughs relative to the ARMA filter. The peak gain is at the high end of the pass band, with a peak just under 2.0, while several points in the pass band have gains near zero.

Panel 6 (right column, bottom row) shows that both filtered series have effectively blocked the high-frequency part of the original spectrum, but that both seem to pass significant amounts at the lowest frequencies. The results shown in Table 4, similar for both models, apparently reflect this low-frequency leakage. The correlation of optimal and ideal filter estimates is only about 75 per cent, and the noise-signal ratio ranges from 72 per cent for the kernel model to 87 per cent for the ARMA model.

### *3.4.1 The reliability of current estimates of the output gap*

Estimates of the reliability of output gaps are of particular interest for the design of optimal monetary policies. It is therefore of interest to compare the above results with other recent estimates in the literature.

Christiano and Fitzgerald (1999) solve roughly the same optimal filter problem solved in this paper.[23] However, on the basis of experiments with models in the IMA(1,q) class, they conclude that most economic time series can be nearly optimally filtered if we assume that they are random walks and solve for the corresponding approximately optimal filter.[24] In that case, the filter weights become functions of only the cut-off frequencies and the sample size. They calculate that the correlation between their nearly optimal filter and the ideally filtered measure of the business cycle is roughly 0.65 and that the noise-signal ratio is 0.77.[25] This is a lower correlation than we find for either model, while their noise-signal ratio is too optimistic relative to ARMA results but slightly pessimistic relative to those of the kernel model.

Orphanides and van Norden (1999, 2002) and Cayen (2001) study other filters that do not have optimal band-pass properties but are nonetheless used to measure business cycles. They compare the rolling estimates produced when such filters are applied at the end-of-sample with historical estimates produced after many subsequent years of data are available.[26] The size of this revision

23. See Appendix A for a discussion.
24. The fact that they use low-order IMA models presumably guarantees that the optimal weights and their approximate weights will differ only for the last few observations, and even then not by very much. It would be interesting to see whether the usefulness of the random walk approximation would be sustained if kernel or ARIMA models were used instead to derive optimal filters.
25. The correlation is taken from the end-point of the graph in the left column, middle row of their Figure 6, while the signal-noise ratio is given on p. 21 in their paper. Note that their band-pass filter is set to pass all cycles with durations from 2 to 8 years, versus the 6 to 32 quarters used here.
26. In the terminology of these papers, these are the Final and the QuasiFinal estimates. The Final–QuasiFinal revision is a better analogue to the estimation error considered in this paper, since both estimates ignore the role of uncertainty in the underlying data-generating process.

in estimated business cycles corresponds to the difference between our ideal and optimal filter estimates. Using U.S. data, Orphanides and van Norden (1999) find correlation coefficients ranging from 0.63 to 0.96, depending on the model used.[27] Cayen (2001) uses Canadian data and finds correlation coefficients for the same models ranging from 0.70 to 0.84.[28]

We can also reconstruct noise-signal ratios for Orphanides and van Norden (1999) based on the ratio of the reported standard deviation of the revisions to the standard deviation of the final measure of the output gap. These figures are reported in Table 5; the values range from 0.34 to 0.79.

While it may appear to be counterintuitive that non-optimal models appear sometimes to give better correlations or noise-signal ratios than the optimal measures developed in this paper, it should be remembered that the two are not strictly comparable, since their definitions of trend and cycle differ. These results, however, suggest that the optimal frequency-based techniques should not be expected to give markedly more accurate estimates than other sophisticated time-series methods.

## 4. Conclusions

The derivations in Appendix A show how to construct optimal band-pass filters for the ARIMA models commonly used with macroeconomic time series. Together with the above-noted applications of such filters, this produces several interesting results.

First, it illustrates that the accuracy of such filters can vary considerably. When a series has little or no predictability, as is the case for productivity growth, this limits our ability to measure the current long-term trend. We found therefore that current estimates of trend productivity growth have correlations barely above 70 per cent with comparable estimates constructed with the benefit of hindsight; put another way, the filters' noise-signal ratio approaches 1. This implies that the measurement of productivity trends is the most difficult of the three problems considered in this paper.

While the results for inflation were somewhat sensitive to the representation chosen, they illustrate how increased predictability improves our measurement of current trends. If we accept that the kernel model is the more plausible representation of inflation dynamics, it implies that current inflation trends can be measured with considerable accuracy, giving 95 per cent correlations with the best ex-post measures and noise-signal ratios barely above 10 per cent.

27. Orphanides and van Norden (1999, Table 1, 32). Results are correlations between QuasiFinal and Final estimates for the Watson, Clark, and Harvey-Jaeger models.
28. Cayen (2001, Table 1, 41).

Persistence acts as a double-edged sword, however, when we seek to measure current cycles, as shown by the results for business cycle measurement. On the one hand, it improves the amount of information available about the future of the series, thereby reducing the difference between current and future estimates of trend. On the other hand, by increasing the relative amount of noise to be filtered out, it increases the potential effects of leakage and therefore of measurement error. In the case of business cycle measurement, we see that optimal filters do not perform especially well; their correlations and noise-signal ratios are not much better than those for productivity growth.

The results for business cycle measurement are somewhat surprising in light of previous work that examined the performance of non-frequency-based models of trends and cycles. Comparison of these results seems to show that the latter will in some cases perform as well as or better than the optimal methods analyzed here. The comparisons are potentially misleading, however, since the definitions of trends and cycles are not comparable across models. The previous work also focused on models in which output was assumed to follow a stochastic trend, rather than the deterministic trend assumed here. A reconciliation of these results should examine the extent to which the results for frequency-based filters are sensitive to the assumption of trend-stationarity.

More generally, the sensitivity of optimal filters to the assumed dynamics of the data series requires further evaluation. This would allow closer scrutiny of Christiano and Fitzgerald's (1999) claim that the assumption of random walk dynamics is adequate for most macroeconomic time series, which is a potentially important simplification for applied work. It would also have implications for the accuracy with which business cycles may realistically be measured. The only source of error considered in this paper's analysis is the extent to which estimated cycles will be revised as new observations become available. As noted in Orphanides and van Norden (1999, 2002), other sources include estimation error in the autocovariance function, data revision, and model misspecification. The results presented herein should therefore be viewed as lower bounds on the total measurement error in frequency-based estimations of current trends and cycles.

## References


Baxter, M. and R. King. 1999. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series." *Review of Economics and Statistics* 81(4): 575–93.

Burns, A.F. and W.C. Mitchell. 1947. *Measuring Business Cycles*. NBER, New York.

Cayen, J.-P. 2001. "Fiabilité des estimations de l'écart de production au canada." *École des Hautes Études Commerciales memoire*.

Christiano, L.C. and T.J. Fitzgerald. 1999. "The Band Pass Filter." Federal Reserve Bank of Cleveland Working Paper No. 99-06. Also published as NBER Working Paper No. 7257.

Cogley, T. and J.M. Nason. 1995. "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research." *Journal of Economic Dynamics and Control* 19: 253–78.

Den Haan, W.J. and S. Sumner. 2001. "The Comovements Between Real Activity and Prices in the G7." University of California San Diego Department of Economics Working Paper No. 2001-05.

Gerlach, S. and F. Smets. 1997. "Output Gaps and Inflation: Unobservable-Components Estimates for the G-7 Countries." Bank for International Settlements, Basel. Photocopy.

Gomez, V. 2001. "The Use of Butterworth Filters for Trend and Cycle Estimation in Economic Time Series." *Journal of Business and Economic Statistics* 19(3): 365–73.

Guay, A. and P. St-Amant. 1997. *Do Mechanical Filters Provide a Good Approximation of Business Cycles?* Technical Report No. 78. Ottawa: Bank of Canada.

Harvey, A.C. 1985. "Trends and Cycles in Macroeconomic Time Series." *Journal of Business and Economic Statistics* 3: 216–27.

Harvey, A.C. and A. Jaeger. 1993. "Detrending, Stylized Facts and the Business Cycle." *Journal of Applied Econometrics* 8: 231–47.

Hodrick, R. and E. Prescott. 1997. "Post-War Business Cycles: An Empirical Investigation." *Journal of Money, Credit, and Banking* 29(1): 1–16. (Working paper version cited as Hodrick and Prescott 1977.)

Kaiser, R. and A. Maravall. 2001. *Measuring Business Cycles in Economic Time Series.* Lecture Notes in Statistics, volume 154. Springer Verlag.

Koopmans, L.H. 1974. *The Spectral Analysis of Time Series*. San Diego, CA: Academic Press.

Kuttner, K.N. 1994. "Estimating Potential Output as a Latent Variable." *Journal of Business and Economic Statistics* 12(3): 361–68.

Morley, J., C. Nelson, and E. Zivot. 1999. *Why are Beveridge-Nelson and Unobserved-Component Decompositions of GDP so Different?* University of Washington. Photocopy.

Orphanides, A. and S. van Norden. 1999. "The Reliability of Output Gap Estimates in Real Time." Finance and Economics Discussion Series, 1999–38, Federal Reserve Board, Washington.

———. 2002. "The UnReliability of Output Gap Estimates in Real Time." *The Review of Economics and Statistics*. Forthcoming.

Ouliaris, S. and P.C.B. Phillips. 1995. *Coint 2.0: GAUSS Procedures for Cointegrated Regressions.* Aptech Systems.

Pedersen, T.M. 1998. "The Hodrick-Prescott Filter, The Slutzky Effect, and the Distortionary Effect of Filters." University of Copenhagen, Institute of Economics Discussion Paper No. 98/09.

Pollock, D.S.G. 2000. "Trend Estimation and De-trending via Rational Square-Wave Filters." *Journal of Econometrics* 99(2): 317–34.

St-Amant, P. and S. van Norden. 1997. *Measurement of the Output Gap: A Discussion of Recent Research at the Bank of Canada*. Technical Report No. 79: Ottawa: Bank of Canada.

Schleicher, C. 2001. "Approximate Bandpass Filters for ARIMA Models." University of British Columbia. Photocopy.

Setterfield, M.A., D.V. Gordon, and L. Osberg. 1992. "Searching for a Will o' the Wisp: An Empirical Study of the NAIRU in Canada." *European Economic Review* 36(1): 119–36.

Staiger, D., J.H. Stock, and M.W. Watson. 1997. "The NAIRU, Unemployment and Monetary Policy." *Journal of Economic Perspectives* 11(1): 33–49.

Stock, J. and M. Watson. 1998. "Business Cycle Fluctuations in U.S. Macroeconomic Time Series." NBER Working Paper No. 6528. Also available in *The Handbook of Macroeconomics*, edited by J.B. Taylor and M. Woodford. Amsterdam: Elsevier Science.

**Table 1: Estimated ARMA Models**

| Series | $Q_t$ | $\pi_t$ | $Y_t$ |
|---|---|---|---|
| # AR parameters | 0 | 0 | 1 |
| $\rho$ | | | 0.968 |
| # MA parameters | 0 | 12 | 1 |
| $\sum\theta_i$ | | 1.46 | 0.312 |
| $\sigma^2 \times 10^6$ | 65.28 | 6.619 | 74.36 |
| # observations | 141 | 631 | 162 |
| Frequency | Quarterly | Monthly | Quarterly |

**Table 2: Optimal Filter for Productivity Growth Trends**

| Statistic[a] | ARMA model | Kernel model |
|---|---|---|
| 1) Variance of raw series (x $10^6$)[b] | 65.28 | 67.15 |
| 2) Variance of ideally filtered series (x $10^6$)[c] | 4.08 | 3.32 |
| 3) Variance of optimally filtered series (x $10^6$)[d] | 2.14 | 1.70 |
| 4) MSE of filtered estimate (x $10^6$)[e] | 1.94 | 1.62 |
| 5) Correlation with ideal estimate[f] | 0.725 | 0.715 |
| 6) N/S ratio[g] | 0.903 | 0.954 |

a. All integrals were numerically approximated as Reimann sum over 10,000 equally spaced frequencies on the interval $(-\pi, \pi)$.

b. Differences between the ARMA and kernel estimates of the variance result from approximations in the construction of the theoretical autocorrelations of the ARMA model; the kernel estimates precisely match the sample variance of the series.

c. Calculated as the integral of the spectral density over the interval $(-u,u)$.

d. Calculated as the integral of the spectral density over the interval $(-\pi, \pi)$.

e. Calculated as (2) - (3).

f. Calculated as $\sqrt{(3)/(2)}$.

g. Calculated as (4) / (3).

**Table 3: Optimal Filter for Inflation Trends[a]**

| Statistic | ARMA model | Kernel model |
|---|---|---|
| 1) Variance of raw series (x $10^6$) | 11.90 | 15.34 |
| 2) Variance of ideally filtered series (x $10^6$) | 1.63 | 6.60 |
| 3) Variance of optimally filtered series (x $10^6$) | 1.04 | 5.94 |
| 4) MSE of filtered estimate (x $10^6$) | 0.60 | 0.66 |
| 5) Correlation with ideal estimate | 0.796 | 0.949 |
| 6) N/S ratio | 0.577 | 0.111 |

a. See footnotes for Table 2.

**Table 4: Optimal Filter for Business Cycles[a]**

| Statistic | ARMA model | Kernel model |
|---|---|---|
| 1) Variance of raw series (x $10^6$) | 2006 | 4524 |
| 2) Variance of ideally filtered series (x $10^6$) | 168.9 | 442.0 |
| 3) Variance of optimally filtered series (x $10^6$) | 90.5 | 257.2 |
| 4) MSE of filtered estimate (x $10^6$) | 78.4 | 184.8 |
| 5) Correlation with ideal estimate | 0.732 | 0.763 |
| 6) N/S ratio | 0.866 | 0.719 |

a. See footnotes for Table 2.

**Table 5: Reconstructed Noise-Signal Ratios from Orphanides and van Norden (1999)**

| Model | Errors[a] (1) | Final[b] (2) | Noise/Signal = (1)/(2) |
|---|---|---|---|
| Clark | 1.11 | 2.11 | 0.53 |
| Harvey-Jaeger | 1.22 | 1.55 | 0.79 |
| Watson | 1.16 | 3.44 | 0.34 |

a. Figures are the reported standard deviations for Final - Quasi-Final revisions, taken from Orphanides and van Norden (1999, Table 4, 36).

b. Figures are the reported standard deviations for Final output gaps, taken from Orphanides and van Norden (1999, Table 1, 32).

**Figure 1: Squared Gain of Band-Pass Filters: Ideal and Approximate**

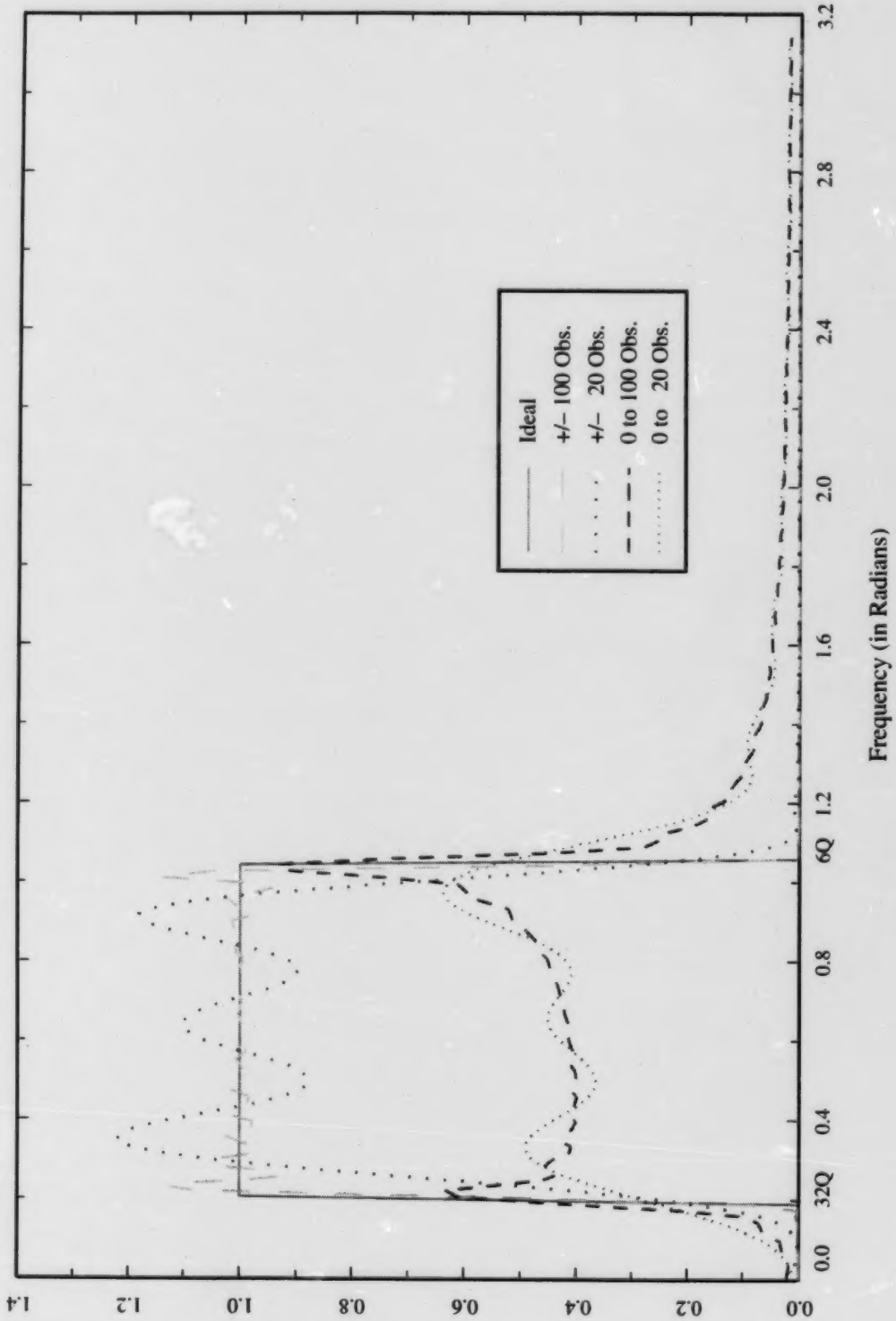


**Figure 2: Filtering Quarterly Productivity Growth for its Current Trend**

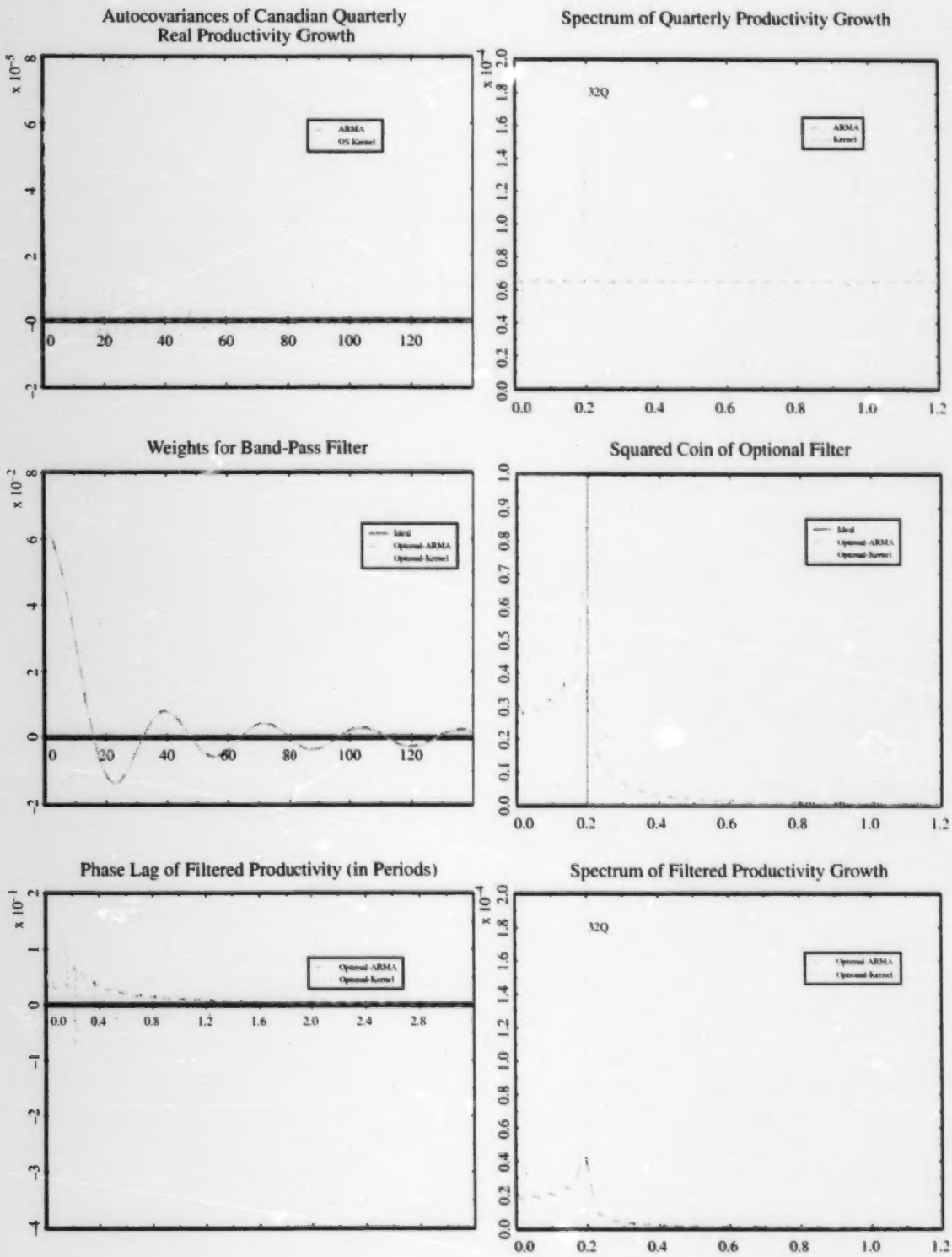


**Figure 3: Filtering Monthly Data on Quarterly Inflation for its Current Trend**

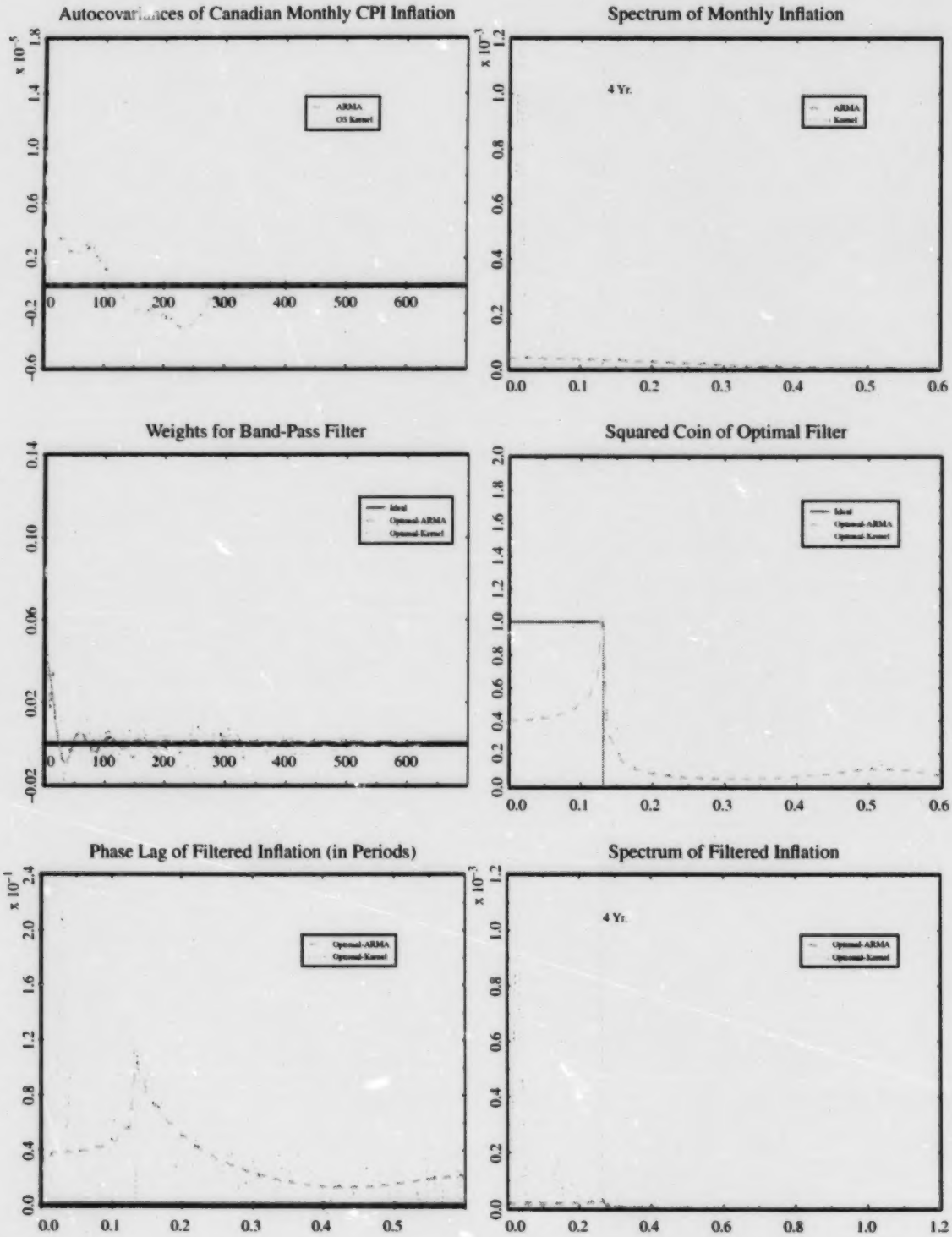
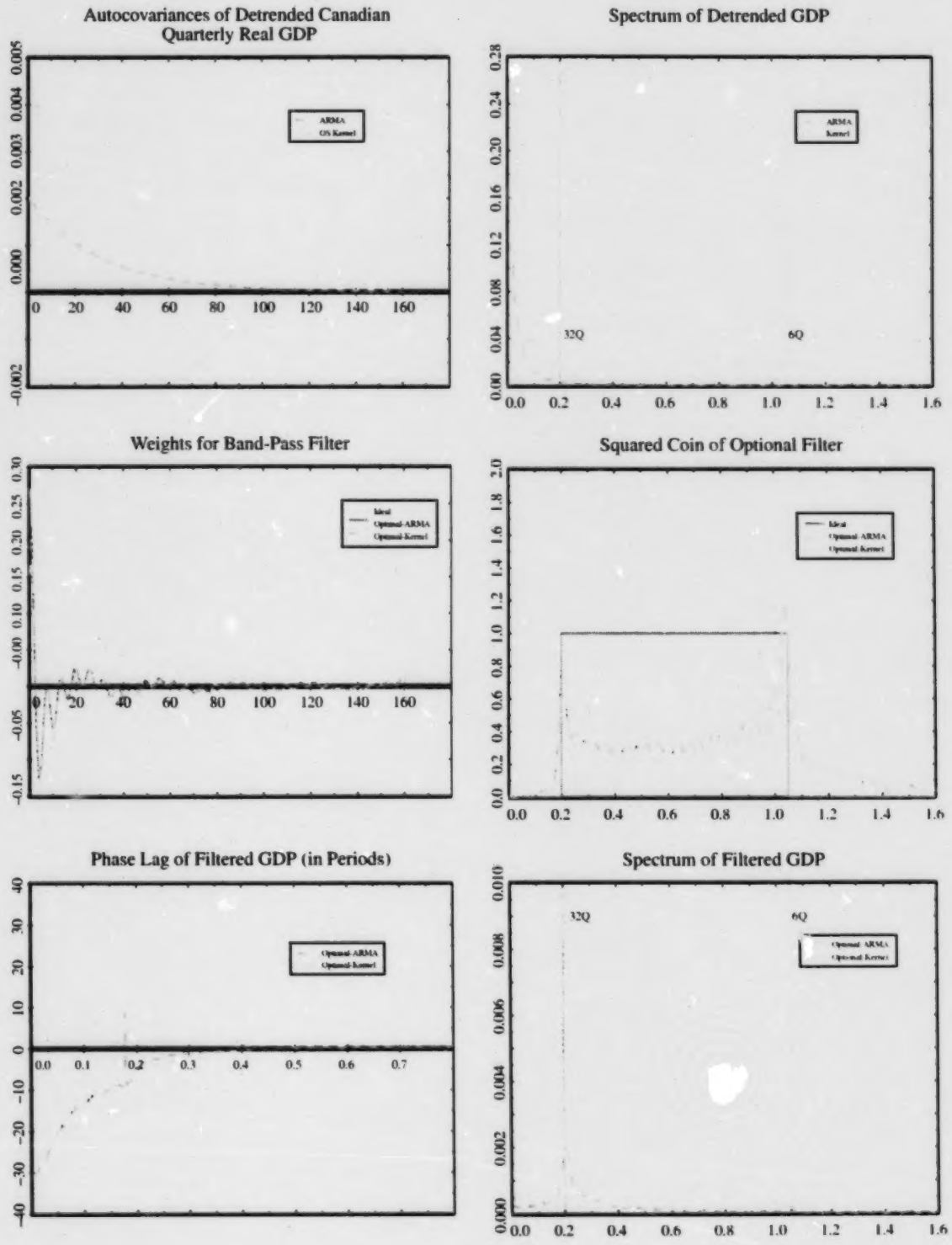


**Figure 4: Filtering Quarterly Output for the Current Business Cycle**



# Appendix A: Derivation of the Optimal Filter

This appendix presents a derivation of the optimal one-sided band-pass filter. It draws on Christiano and Fitzgerald (1999), who derive the optimal filter estimated at each point in the sample for an IMA(1,q) process.The proof is given in two stages. In section A.2 we derive the solution under the assumption that our process is a stationary MA($q$) process with finite variance. Sections A.3 and A.4 discuss extensions to the ARMA and ARIMA cases. Appendix B then compares the contribution of Christiano and Fitzgerald (1999) with that of Koopmans (1974), and provides additional intuition about the nature of the solution.

After this appendix was prepared, Schleicher (2001) provided a more compact derivation of the optimal filter for the ARIMA case, which extends the results below to any arbitrary point in the sample, instead of just the endpoint.

## A.1 Notation and Basic Assumptions

We have a sample of $T$ discrete observations on some stochastic series $\{y_t\}$. We assume that $\{y_t\}$ contains no deterministic components (constants or non-stochastic trends).[1] For the remainder of this section, we also assume that $\{y_t\}$ is covariance stationary with a known MA($q$) representation for finite $q$. We also require that the covariance matrix of $[y_t, y_{t-1}, y_{t-2}, \ldots, y_{t-T+1}]$ exists and has a unique inverse. Finally, we will assume that the power spectrum of y also exists and is given by[2]

$$f_y(\omega) = \frac{1}{2\pi} \cdot \sum_{k=-\infty}^{\infty} \sigma_y(k) \cdot e^{-i\omega k}, \tag{A.1}$$

where $\sigma_y(k)$ is the covariance between $y_t$ and $y_{t-k}$.

Ideally, we would want to choose $\{x_t, c_t\}$ to partition the spectrum of $y_t$, so that $c_t$ contains all the fluctuations in $y_t$ with frequencies between some lower limit, $l$, and some upper limit, $u$, $0 < l < u \leq \pi$, and $x_t$ contains those fluctuations with frequencies below $l$ or above $u$.[3] We can do so with an infinite-order time-invariant linear filter,

---

1. Typically, one would regress the raw data series on a polynomial time trend and use the residuals as $y_t$, which may be stationary or stochastically integrated. As in all spectral analysis, we ignore the potentially important effects of any imperfections in this detrending.
2. Covariance stationarity plus finite variance is sufficient to guarantee the properties on the covariance matrix. The power spectrum conditions will be satisfied if the coefficients on the MA representation are absolutely summable; covariance stationarity guarantees that they are square summable.
3. For example, isolating the fluctuations between 6 and 32 quarters in length with quarterly data would correspond to $l = \pi/16$ and $u = \pi/3$.

$$B(L) \equiv \sum_{j=-\infty}^{\infty} B_j \cdot L^j , \tag{A.2}$$

such that $c_t = B(L) \cdot y_t$ and $x_t = y_t - c_t = y_t \cdot (1 - B(L))$, where $L$ is the lag operator. The coefficients of this ideal filter are given by[4]

$$\begin{aligned} B_j &= \frac{\sin ju - \sin jl}{\pi j} \text{ for } j \neq 1 \\ &= \frac{u-l}{\pi} \text{ otherwise.} \end{aligned} \tag{A.3}$$

Our problem is to choose a set of filter weights, $\{\hat{B}_j\}$, to solve

$$\min_{\{\hat{B}_j\}} \quad \int_{-\pi}^{\pi} \left| B(e^{-i\omega}) - \hat{B}(e^{-i\omega}) \right|^2 \cdot f_y(\omega) d\omega, \tag{A.4}$$

where

$$\hat{B}(z) = \sum_{j=0}^{T-1} \hat{B}_j \cdot z^j , \tag{A.5}$$

for complex $z$ on the unit circle.

## A.2 Solution for the MA($q$) Case

### A.2.1 First-order conditions

Rewrite equation (A.4) as

$$\min_{\{\hat{B}_j\}} \quad \int_{-\pi}^{\pi} \delta(\omega) \cdot \delta(-\omega) \cdot f_y(\omega) d\omega \qquad \text{where } \delta(\omega) = B(e^{-i\omega}) - \hat{B}(e^{-i\omega}) . \tag{A.6}$$

Differentiating (A.4) with respect to each element of $\{\hat{B}_j\}$, we obtain $T$ first-order conditions

4. Christiano and Fitzgerald (1999) further impose the restriction that $\sum_j B_j = 0$, which is reasonable except in the special case of low-pass filters.

$$0 = \int_{-\pi}^{\pi} \left( \delta(\omega) \cdot \frac{\partial}{\partial \hat{B}_j} \delta(-\omega) + \delta(-\omega) \cdot \frac{\partial}{\partial \hat{B}_j} \delta(\omega) \right) \cdot f_y(\omega) d\omega \qquad \forall j = 0, \ldots, T-1. \tag{A.7}$$

In fact,

$$\frac{\partial}{\partial \hat{B}_j} \delta(-\omega) = -e^{i\omega j} \text{ and } \frac{\partial}{\partial \hat{B}_j} \delta(\omega) = -e^{-i\omega j}, \tag{A.8}$$

so (A.7) becomes

$$0 = \int_{-\pi}^{\pi} (\delta(\omega) \cdot -e^{i\omega j} + \delta(-\omega) \cdot -e^{-i\omega j}) \cdot f_y(\omega) d\omega$$

$$\therefore \int_{-\pi}^{\pi} [B(e^{-i\omega}) \cdot e^{i\omega j} + B(e^{i\omega}) \cdot e^{-i\omega j}] \cdot f_y(\omega) d\omega = \tag{A.9}$$

$$\int_{-\pi}^{\pi} [\hat{B}(e^{-i\omega}) \cdot e^{i\omega j} + \hat{B}(e^{i\omega}) \cdot e^{-i\omega j}] \cdot f_y(\omega) d\omega \qquad \forall j = 0, \ldots, T-1.$$

Noting that $f_{\Delta y}(\omega)$ is symmetric about 0, we obtain

$$\int_{0}^{\pi} [B(e^{-i\omega}) \cdot e^{i\omega j} + B(e^{i\omega}) \cdot e^{-i\omega j}] \cdot f_y(\omega) d\omega = \tag{A.10}$$

$$\int_{-\pi}^{0} [B(e^{-i\omega}) \cdot e^{i\omega j} + B(e^{i\omega}) \cdot e^{-i\omega j}] \cdot f_y(\omega) d\omega.$$

and

$$\int_{0}^{\pi} [\hat{B}(e^{-i\omega}) \cdot e^{i\omega j} + \hat{B}(e^{i\omega}) \cdot e^{-i\omega j}] \cdot f_y(\omega) d\omega = \tag{A.11}$$

$$\int_{-\pi}^{0} [\hat{B}(e^{-i\omega}) \cdot e^{i\omega j} + \hat{B}(e^{i\omega}) \cdot e^{-i\omega j}] \cdot f_y(\omega) d\omega.$$

Equation (A.9) implies that

$$\int_{0}^{\pi} [B(e^{-i\omega}) \cdot e^{i\omega j} + B(e^{i\omega}) \cdot e^{-i\omega j}] \cdot f_y(\omega) d\omega = \int_{0}^{\pi} [\hat{B}(e^{-i\omega}) \cdot e^{i\omega j} + \hat{B}(e^{i\omega}) \cdot e^{-i\omega j}] \cdot f_y(\omega) d\omega,$$

$$\therefore \int_{-\pi}^{\pi} B(e^{-i\omega}) \cdot e^{i\omega j} \cdot f_y(\omega) d\omega = \int_{-\pi}^{\pi} \hat{B}(e^{-i\omega}) \cdot e^{i\omega j} \cdot f_y(\omega) d\omega \quad \forall j = 0, \ldots, T-1. \tag{A.12}$$

Using the fact that $\int_{-\pi}^{\pi} e^{i\omega j} d\omega = 2\pi$ when $j = 0$ and $\int_{-\pi}^{\pi} e^{i\omega j} d\omega = 0$ when $j = 1, 2, 3, \ldots$, we can see that (A.12) simply picks out one term at a time from the convolutions of the filter coefficients and the spectrum of $y$.

Because $\sigma_y(k)$ has the properties $\sigma_y(k) = \sigma_y(-k)$ and $\sigma_y(k) = 0 \ \forall k > q$, we can write

$$\int_{-\pi}^{\pi} B(e^{-i\omega}) \cdot e^{i\omega j} \cdot f_y(\omega) d\omega = 2\pi \cdot \sum_{k=-q}^{q} B_{j+k} \cdot \sigma_y(k) = 2\pi \cdot \vec{B}^{j\prime} \cdot \vec{\sigma}_y, \tag{A.13}$$

where

$\vec{B}^j = [B_{j-q}, B_{j-q+1}, B_{j-q+2}, \ldots, B_{j+q}]'$, a $2q+1$ column vector, and

$\vec{\sigma}_y = [\sigma_y(-q), \sigma_y(-q+1), \sigma_y(-q+2), \ldots, \sigma_y(q)]'$, a $2q+1$ column vector.

We can similarly transform the right-hand side of (A.12), taking care to account for the truncation of $\hat{B}(e^{-i\omega})$, and obtain

$$\int_{-\pi}^{\pi} \hat{B}(e^{-i\omega}) \cdot e^{i\omega j} \cdot f_y(\omega) d\omega = 2\pi \cdot \sum_{k=-j}^{T-1-j} \hat{B}_{j+k} \cdot \sigma_y(k) = 2\pi \cdot \vec{\beta}' \cdot \vec{\sigma}_{y,j} \tag{A.14}$$

where

$\vec{\beta} = [\hat{B}_0, \ldots, \hat{B}_{T-1}]'$, a length $T$ column vector, and

$\vec{\sigma}_{y,j} = [\sigma_y(-j), \sigma_y(-j+1), \sigma_y(-j+2), \ldots, \sigma_y(T-1-j)]'$, a length $T$ column vector.

This allows us to rewrite (A.12) compactly as

$$\vec{B}^{j\prime} \cdot \vec{\sigma}_y = \vec{\sigma}_{y,j}' \cdot \vec{\beta} \quad \forall j = 0, \ldots, T-1. \tag{A.15}$$

### A.2.2 Solving the system of first-order conditions

In (A.15), $\vec{\sigma}_y$ and $\vec{\sigma}_j$ are known and determined by the data, while $\vec{B}^j$ is given by (3) in the main text. This leaves us with $T$ linear equations to solve for the $T$ elements of $\vec{\beta}$. Stacking these $T$ equations gives us

$$\hat{B} \cdot \hat{\sigma}_y = \hat{\Sigma}_y \cdot \hat{\beta}, \tag{A.16}$$

where

$\hat{B} = [\hat{B}^0, \hat{B}^1, \hat{B}^2, \ldots, \hat{B}^{T-1}]'$, a $T \times 2q+1$ matrix, and

$\hat{\Sigma}_y = [\hat{\sigma}_{y,0}, \hat{\sigma}_{y,1}, \hat{\sigma}_{y,2}, \ldots, \hat{\sigma}_{y,T-1}]'$, a $T \times T$ matrix.

The solution is

$$\hat{\beta} = \hat{\Sigma}_y^{-1} \cdot \hat{B} \cdot \hat{\sigma}_y, \tag{A.17}$$

provided that $\hat{\Sigma}_y^{-1}$ exists and is unique. We know that

$$\hat{\Sigma}_y = \begin{bmatrix} \sigma_y(0) & \sigma_y(1) & \sigma_y(2) & \ldots & \sigma_y(T-1) \\ \sigma_y(-1) & \sigma_y(0) & \sigma_y(1) & \ldots & \sigma_y(T-2) \\ \sigma_y(-2) & \sigma_y(-1) & \sigma_y(0) & \ldots & \sigma_y(T-3) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \sigma_y(-T+1) & \sigma_y(-T+2) & \sigma_y(-T+3) & \ldots & \sigma_y(0) \end{bmatrix} \tag{A.18}$$

$$= \begin{bmatrix} \sigma_y(0) & \sigma_y(1) & \sigma_y(2) & \ldots & \sigma_y(T-1) \\ \sigma_y(1) & \sigma_y(0) & \sigma_y(1) & \ldots & \sigma_y(T-2) \\ \sigma_y(2) & \sigma_y(1) & \sigma_y(0) & \ldots & \sigma_y(T-3) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \sigma_y(T-1) & \sigma_y(T-2) & \sigma_y(T-3) & \ldots & \sigma_y(0) \end{bmatrix},$$

which is therefore just the covariance matrix of the vector $[y_t, y_{t-1}, y_{t-2}, \ldots, y_{t-T+1}]$. Our solution is therefore applicable whenever this covariance matrix exists and is of full rank.

## A.3 The Stationary ARMA(p,q) Case

If we wish to extend our analysis from the MA($q$) to the ARMA($p$,$q$) case, we can use the Wold Representation Theorem to recast the ARMA model as an infinite-order MA. Assuming that the other conditions mentioned before (A.1) are respected, we can continue to write our optimization problem as before and derive the same first-order conditions. The analysis then proceeds until (A.13), which becomes

$$\int_{-\pi}^{\pi} B(e^{-i\omega}) \cdot e^{i\omega j} \cdot f_y(\omega) d\omega = 2\pi \cdot \sum_{k=-\infty}^{\infty} B_{j+k} \cdot \sigma_y(k) = 2\pi \cdot \vec{B}^{j\prime} \cdot \vec{\sigma}_y, \qquad \text{(A.19)}$$

where $\vec{B}^j$ and $\vec{\sigma}_y$ are now doubly infinite-dimensional vectors. We require only that their dot-product is well defined and finite. Stacking our $T$ equations in (A.16) proceeds as before, the only change being the dimensions of $\vec{B}$ and $\vec{\sigma}_y$ but not their product. The final solution in equation (5) in the main text is therefore unchanged.

Applying the filter to ARMA processes will in practice require that we truncate the infinite sum at some point, hoping that the omitted covariance terms are sufficiently close to zero. Of course, for given parameters for the ARMA process, theoretical autocovariances may easily be calculated, so that the only limitations to the precision of our calculations are computing power and storage capacity. It would be unrealistic, however, to believe that we can infer much about autocovariances beyond the $N$-$1$th lag from a sample of only $N$ observations.

## A.4 The ARIMA(p,d,q) Case

We can try to analyze non-stationary ARIMA(p,d,q) processes by first noting that, after differencing $d$ times, we have ARMA(p,q) processes. It would therefore be tempting to simply assume a sample of size $T+d$, difference it $d$ times, and then apply the analysis as for stationary ARMA models to the differenced data. There are two faults with this approach, both of which can be seen in our objective function (A.4). First, instead of attempting to match the ideal band-pass filter, $B(L)$, on the non-stationary data, we are trying to match it on the differenced data, which is equivalent to setting $B(L) \cdot (1-L)^d$ as our target filter. Second, the power spectrum used to weight the deviations from the ideal filter would now be $f_{\Delta^d y}(\omega)$ instead of $f_y(\omega)$; it is not clear how to motivate such a choice in practice.

Instead, note that equation (3) in the main text implies that we can factor

$$B(L) = (1-L) \cdot b(L) = (1-L)^2 \cdot bb(L), \qquad \text{(A.20)}$$

$$b(L) = \sum_{j=-\infty}^{\infty} b_j \cdot L^j \text{ and } bb(L) = \sum_{j=-\infty}^{\infty} bb_j \cdot L^j. \qquad \text{(A.21)}$$

This implies that the ideal band-pass filter will render stationary any series integrated of order no more than 2.[5] From our objective function in (A.4), we can see that its value will be less than

5. See Den Haan and Sumner (2001) for a related discussion.

infinity if and only if our one-sided filter, $\hat{B}(L)$, also renders $y_t$ stationary. We can therefore presume that the optimal one-sided filter may similarly be factored as

$$\hat{B}(L) = (1-L)\cdot\hat{b}(L) = (1-L)^2\cdot\hat{bb}(L), \tag{A.22}$$

$$\hat{b}(L) = \sum_{j=-\infty}^{\infty}\hat{b}_j\cdot L^j \text{ and } \hat{bb}(L) = \sum_{j=-\infty}^{\infty}\hat{bb}_j\cdot L^j. \tag{A.23}$$

This common factoring of the ideal and one-sided filters allows us to find solutions in the cases of I(1) and I(2) processes. In section A.4.1, we detail the proof for the case of IMA(1,q) processes. We then briefly discuss the IMA(2,q) case. Extensions to the ARIMA case should be clear from the above discussion of the ARMA case.

### A.4.1 The IMA(1,q) case

Suppose that we have $T+1$ observations on a process $y_t \sim IMA(1, q)$. Our minimization problem in (A.6) can then be rewritten as

$$\min_{\{\hat{b}_j\}} \int_{-\pi}^{\pi}\delta(\omega)\cdot(1-e^{-i\omega})\cdot\delta(-\omega)\cdot(1-e^{i\omega})\cdot f_y(\omega)d\omega, \tag{A.24}$$

where now $\delta(\omega) = b(e^{-i\omega}) - \hat{b}(e^{-i\omega})$. We can also now factor

$$f_y(\omega) = f_{\Delta y}(\omega)\cdot(1-e^{-i\omega})^{-1}\cdot(1-e^{i\omega})^{-1}, \tag{A.25}$$

where $f_{\Delta y}(\omega)$ is the power spectrum for the first-difference of $y_t$. This means that (A.24) simplifies to

$$\min_{\{\hat{b}_j\}} \int_{-\pi}^{\pi}\delta(\omega)\cdot\delta(-\omega)\cdot f_{\Delta y}(\omega)d\omega. \tag{A.26}$$

We now have an optimization problem that is analogous to our original problem in (A.4). The only differences are three substitutions:

(i) we use the power spectrum of $\Delta y_t$ rather than of $y_t$,

(ii) the function we seek to match is now $b(L)$ rather than $B(L)$, and

(iii) the optimization is over the $T$ coefficients of the differenced filter, $\hat{b}_j$, rather than the $T+1$ coefficients, $\hat{B}_j$.

As in the original proof, we work with the spectrum of a stationary series (now $\Delta y_t$) on which we have $T$ observations. We therefore arrive at an analogous first-order condition, replacing (A.15) with

$$\vec{\hat{b}}^{j\prime} \cdot \vec{\sigma}_{\Delta y} = \vec{\sigma}_{\Delta y,\, j}{}' \cdot \vec{\hat{\beta}}_1 \quad \forall j = 0, \ldots, T-1, \tag{A.27}$$

where

$\vec{\hat{b}}^j = [b_{j-q}, b_{j-q+1}, b_{j-q+2}, \ldots, b_{j+q}]'$, a $2q+1$ column vector,

$\vec{\sigma}_{\Delta y} = [\sigma_{\Delta y}(-q), \sigma_{\Delta y}(-q+1), \sigma_{\Delta y}(-q+2), \ldots, \sigma_{\Delta y}(q)]'$, a $2q+1$ column vector,

$\sigma_{\Delta y}(k)$ is the autocovariance function for $\Delta y_t$,

$\vec{\sigma}_{\Delta y,\, j} = [\sigma_{\Delta y}(-j), \sigma_{\Delta y}(-j+1), \sigma_{\Delta y}(-j+2), \ldots, \sigma_{\Delta y}(T-1-j)]'$, a length $T$ column vector, and

$\vec{\hat{\beta}}_1 = [\hat{b}_0, \ldots, \hat{b}_{T-1}]'$, a length $T$ column vector.

These may be stacked and solved for the optimal $b(L)$, given by

$$\vec{\hat{\beta}}_1 = \hat{\Sigma}_{\Delta y}^{-1} \cdot \hat{b} \cdot \vec{\sigma}_{\Delta y}, \tag{A.28}$$

where

$\hat{b} = [\vec{\hat{b}}^0, \vec{\hat{b}}^1, \vec{\hat{b}}^2, \ldots, \vec{\hat{b}}^{T-1}]'$, a $T$ x $2q+1$ matrix,

$\hat{\Sigma}_{\Delta y} = [\vec{\sigma}_{\Delta y,\, 0}, \vec{\sigma}_{\Delta y,\, 1}, \vec{\sigma}_{\Delta y,\, 2}, \ldots, \vec{\sigma}_{\Delta y,\, T-1}]'$, a $T$ x $T$ matrix.

While (A.28) determines $\vec{\hat{\beta}}_1$, this gives us only $T$ conditions with which to identify the $T+1$ coefficients, $\hat{B}_j$. The remaining condition is (A.22), which implies that the sum over $j$ of $\hat{B}_j$ must equal zero.

### A.4.2 The IMA(2,q) case

The proof for the IMA(2,q) case with $T+2$ observations proceeds analogously to the IMA(1,q) case; we again transform the original minimization problem into one involving the power

spectrum of a stationary series. This means working throughout the proof with $T$ observations on $\Delta^2 y$, its power spectrum, $f_{\Delta^2 y}(\omega)$, and autocovariance function, $\sigma_{\Delta^2 y}(k)$, the second difference of the ideal band-pass filter, $bb(L)$, and recovering the optimal one-sided filter for second-differenced data, $\hat{bb}(L)$.

The basic intuition should by now be clear. Because we know that the ideal two-sided filter will give us a stationary series, we can transform the optimization problem into an equivalent problem using suitably differenced data and a suitably differenced ideal filter. This approach may break down if we try to go beyond the I(2) case, since the ideal filter is no longer sure to give a stationary filtered series. Fortunately, the vast majority of economic time series do not require models with orders of integration larger than 2.[6]

6. An alternative approach would be to change the definition of the ideal filter to ensure stationarity.

# Appendix B: Historical Antecedents

The tools of frequency-domain analysis and signal processing theory have been available for over a quarter-century. The idea of deriving an optimal filter defined as in equation (A.4) is an obvious one, and the derivation of its closed-form solution is not difficult. Despite this, it is barely mentioned or discussed in modern econometrics or macroeconomics.[7] Christiano and Fitzgerald (1999) derive the optimal filtered estimated at each point in the sample for an IMA(1,q) process.[8] They make no specific references to previous derivations of this type. They refer, however, to Stock and Watson's (1998) time-domain procedure of forecasting future values of output to "pad" the sample and applying the Baxter-King filter, but they do not clearly state its relationship.[9]

Koopmans (1974), however, gives the closed-form solution for a closely related problem.[10] He considers the case of two weakly stationary processes $\{x_t, y_t\}$, where we observe only $y_t$ but we know their cross-spectral density, $f_{xy}(\omega)$, as well as their spectral densities, $f_x(\omega)$, $f_y(\omega)$. (Recall that knowing $f_{xy}(\omega)$ is equivalent to knowing their cross-covariance function $C_{xy}(k)$.) We wish to construct the linear filter, $B_v(L)$, which minimizes

$$E([x_{t+v} - B_v(L) \cdot y_t]^2), \tag{B.1}$$

subject to the restriction that $B_v(L)$ is a polynomial in only non-negative powers of $L$.

Koopmans' proof relies on an underlying intuition based on projections. The optimal one-sided filter is simply the projection from the space spanned by $x_{t+v}$ into the space spanned by past and present values of $y_t$, $H_y^-(t)$. This can be broken down into two projection steps, the first being the projection of $x_t$ onto the space spanned by past, present, and future values of $y_t$ ($H_y$), and the second being the projection of that result onto $H_y^-(t)$.[11]

---

7. One reason may be that most introductory treatments of such techniques assume (i) that the data being analyzed are stationary, and (ii) that our data series are very long. Another reason may be that economists are confused about how to apply such techniques to non-stationary data; see Den Haan and Sumner's (2001, section 3.2, especially 11–12) critique of Harvey and Jaeger (1993) and Cogley and Nason (1995).
8. They assert that the proof could be extended to ARIMA (p,1,q), but that this would be "tedious." Schleicher (2001) provides just such an extension, with a proof that is considerably more elegant than the original proof in Christiano and Fitzgerald (1999).
9. This is the method used in the BPFILTER.SRC routine provided by Estima for RATS; the routine is Taylor's translation of Watson's code for Stock and Watson (1998).
10. See Koopmans (1974), particularly section 7.6, 249–52, and section 5.5, especially example 5.5, 147–48. Again, there is no mention of the origin of the derivation, presumably because it was considered to be too well-known or trivial a result.
11. The projection may be broken into two steps because $H_y^-(t)$ is a subspace of $H_y$.

Koopmans begins by noting that the optimal *two-sided* filter has a transfer function given by

$$B(\omega) = f_{xy}(\omega)/f_{y}(\omega). \tag{B.2}$$

Applying the optimal one-sided filter, $B_V(L)$, to our available series, $\{y_1, \ldots, y_T\}$, will give the same result as applying the doubly infinite optimal two-sided filter, $B(L)$, to a doubly infinite "padded" series, $\{\tilde{y}_t\}$, where $\tilde{y}_t = y_t$ for $t = 1, \ldots, T$, $\tilde{y}_t = \hat{y}_t$ otherwise. $\hat{y}_t$ is simply the optimal forecast/backcast value of $y_t$ given the available sample from 1 to T. Stock and Watson's (1998) approach is therefore an approximation of this, which truncates the forecast/backcast at some finite values.

The relationship to the problem at hand can be understood by considering the latent variable, $x_t$. Ideally, we would like to estimate $x_t$ = the (unobserved) output gap. This is possible if we know $C_{xy}(k)$, its autocovariance function with observed output. This will be the case if we postulate a structural model, as in the structural unobserved components used by Harvey (1985), Kuttner (1994), Gerlach and Smets (1997), and Kaiser and Maravall (2001). The optimal estimates we obtain from the above correspond with the solutions these authors obtain by using the standard recursive Wiener-Kolmogorov filtering and smoothing equations. Of course, results may be highly sensitive to the specification of the structural model, as noted by Morley, Nelson, and Zivot (1999).

The band-filtering approach remains agnostic about the structural model by choosing $x_t$ to be the optimally band-filtered component of observed output. Since the form of the optimal filter, $B(L)$, is known, $x_t = B(L) \cdot y_t$ and $f_{xy}(\omega) = B(\omega) \cdot f_{y}(\omega)$. We can therefore construct the optimal finite-sample approximation of $x_t$ as a function of only the dynamic properties of output. In doing so, however, we ignore the fact that this definition of $x_t$ is only an approximation of the output gap that we ultimately care about. To see this, note that as our sample becomes large, our estimation error for $x_t$ in the middle of our sample tends asymptotically towards zero in this approach, whereas in the unobserved approach we have asymptotic standard errors associated with our smoothed estimates.

## B.1 Koopmans' proof

Since $y_t$ is weakly stationary, it has a Wold representation

$$y_t = \sum_{j=0}^{\infty} c_j \cdot \xi_{t-j}, \tag{B.3}$$

where $\xi_t$ is white noise and we define $C(\omega)$ to the transfer function associated with the above filter. The spectrum of $y_t$ is therefore $f_y(\omega) = C(\omega) \cdot f_\xi(\omega)$. If $B(\omega)$ is (as above) the transfer function of the optimal symmetric band-pass filter, then since we seek to estimate $x_{t+v}$, we need to replace it with $B(\omega) \cdot e^{i\omega v}$ to account for the offset of $v$ periods. It therefore follows that the optimal two-sided estimates of $x_t$ are given by

$$\hat{x}_v(t) = \int_{-\pi}^{\pi} e^{i\omega t + v} \cdot B(\omega) \cdot C(\omega) \cdot f_\xi(\omega) = \int_{-\pi}^{\pi} D(\omega) \cdot f_\xi(\omega), \tag{B.4}$$

where

$$D(\omega) = \sum_{-\infty}^{\infty} d_j \cdot e^{i\omega j}.$$

The one-sided problem simply replaces future values of $\xi_t$ with their expectation, which is zero. This means that the optimal one-sided estimate is given by

$$\tilde{x}_v(t) = \int_{-\pi}^{\pi} D_v(\omega) \cdot f_\xi(\omega), \tag{B.5}$$

$$\text{where } D_v(\omega) = \sum_{o}^{\infty} d_j \cdot e^{i\omega j}. \tag{B.6}$$

This derivation has its counterpart in the derivation for the MA(q) presented in Appendix A, section A.2. Equation (5) in the main text shows us that we are projecting the optimal filter $\hat{B}$ into the space spanned by our $T$ observations on $y_t$ via the $\hat{\sigma}_y$ and $\hat{\Sigma}_y$ terms. Note that there are effectively two distinct truncations occurring: one owing to the limited persistence of the underlying MA process and the other owing to the finite sample size.[12]

12. The former is easy to overlook in Koopmans' treatment, as he assumes a possibly infinite MA representation.

## Bank of Canada Working Papers
## *Documents de travail de la Banque du Canada*

**Working papers are generally published in the language of the author, with an abstract in both official languages.** ***Les documents de travail sont publiés généralement dans la langue utilisée par les auteurs; ils sont cependant précédés d'un résumé bilingue.***

### 2002

| | | |
|---|---|---|
| 2002-27 | Habit Formation and the Persistence of Monetary Shocks | H. Bouakez, E. Cardia, and F.J. Ruge-Murcia |
| 2002-26 | Nominal Rigidity, Desired Markup Variations, and Real Exchange Rate Persistence | H. Bouakez |
| 2002-25 | Nominal Rigidities and Monetary Policy in Canada Since 1981 | A. Dib |
| 2002-24 | Financial Structure and Economic Growth: A Non-Technical Survey | V. Dolar and C. Meh |
| 2002-23 | How to Improve Inflation Targeting at the Bank of Canada | N. Rowe |
| 2002-22 | The Usefulness of Consumer Confidence Indexes in the United States | B. Desroches and M-A. Gosselin |
| 2002-21 | Entrepreneurial Risk, Credit Constraints, and the Corporate Income Tax: A Quantitative Exploration | C.A. Meh |
| 2002-20 | Evaluating the Quarterly Projection Model: A Preliminary Investigation | R. Amano, K. McPhail, H. Pioro, and A. Rennison |
| 2002-19 | Estimates of the Sticky-Information Phillips Curve for the United States, Canada, and the United Kingdom | H. Khan and Z. Zhu |
| 2002-18 | Estimated DGE Models and Forecasting Accuracy: A Preliminary Investigation with Canadian Data | K. Moran and V. Dolar |
| 2002-17 | Does Exchange Rate Policy Matter for Growth? | J. Bailliu, R. Lafrance, and J.-F. Perrault |
| 2002-16 | A Market Microstructure Analysis of Foreign Exchange Intervention in Canada | C. D'Souza |
| 2002-15 | Corporate Bond Spreads and the Business Cycle | Z. Zhang |
| 2002-14 | Entrepreneurship, Inequality, and Taxation | C.A. Meh |
| 2002-13 | Towards a More Complete Debt Strategy Simulation Framework | D.J. Bolder |
| 2002-12 | Modelling Financial Instability: A Survey of the Literature | A. Lai |

Copies and a complete list of working papers are available from:
*Pour obtenir des exemplaires et une liste complète des documents de travail, prière de s'adresser à :*


Publications Distribution, Bank of Canada
234 Wellington Street, Ottawa, Ontario K1A 0G9
E-mail: publications@bankofcanada.ca
Web site: http://www.bankofcanada.ca

Diffusion des publications, Banque du Canada
234, rue Wellington, Ottawa (Ontario) K1A 0G9
Adresse électronique : publications@banqueducanada.ca
Site Web : http://www.banqueducanada.ca